İkili ve Çoklu Karşılaştırmalar

İstatistiksel Deney Tasarımı

Birdal Şenoğlu
Şükrü Acıtaş

İçindekiler

Bu bölümde, (2.1) modelinde,

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_a = \mu \tag{1}$$

şeklinde ifade edilen sıfır hipotezinin reddedilmesi durumunda, deneme ortalamaları arasındaki farklılığın hangi deneme veya denemelerden kaynaklandığını belirlemek amacıyla literatürde yaygın olarak kullanılan bazı ikili ve çoklu karşılaştırma metotları tanıtılmıştır.

Fisher (1935) tarafından önerilen **en küçük anlamlı fark** (least significant difference-LSD) metodunda

$$H_0 : \mu_i = \mu_j, \quad \forall i \neq j \tag{2}$$

hipotezini sınamak için

LSD Test İstatistiği

$$LSD = \begin{cases} t_{\alpha/2, df_{\mathbf{Hata}}} \sqrt{\dfrac{2MS_{\mathbf{Hata}}}{n}}, & n_1 = n_2 = \cdots = n_a = n \text{ ise} \\ t_{\alpha/2, df_{\mathbf{Hata}}} \sqrt{MS_{Hata} \left(\frac{1}{n_i} + \frac{1}{n_j} \right)}, & \text{en az biri farklı ise} \end{cases} \tag{3}$$

değeri hesaplanır.

Eğer,

$$|\bar{y}_{i\cdot} - \bar{y}_{j\cdot}| > LSD \tag{4}$$

ise H_0 hipotezi reddedilir. Bir başka deyişle, "*i−inci ve j−inci deneme ortalamaları arasında anlamlı bir farklılık vardır*", denir, bkz Montgomery (2001).

Tukey metodu

$$q = \frac{Maksimum(\bar{y}_{i\cdot}) - Minimum(\bar{y}_{i\cdot})}{\sqrt{\dfrac{MS_{Hata}}{n}}}, \quad i = 1, 2, \cdots, a \tag{5}$$

olarak tanımlanan **Studentlaştırılmış aralık istatistiğine** (Studentized range statistics) dayanır, bkz Kuehl (2000).

Tukey metodunda, denemelerdeki gözlem sayılarının eşit olması halinde $(n_1 = n_2 = \cdots = n_a = n)$

$$H_0 : \mu_i = \mu_j, \quad \forall i \neq j \tag{6}$$

hipotezini sınamak için

HSD Test İstatistiği

$$HSD = q_\alpha(a, df_{Hata})\sqrt{\frac{MS_{Hata}}{n}} \tag{7}$$

değeri hesaplanır.

- Burada, $q_\alpha(a, df_{Hata})$, studentlaştırılmış aralık istatistiğinin tablo değeridir, bkz. May (1952).
- Eğer,

$$|\bar{y}_{i\cdot} - \bar{y}_{j\cdot}| > HSD, \quad \forall i \neq j \tag{8}$$

ise H_0 hipotezi reddedilir. Bir başka deyişle, *"i−inci ve j−inci deneme ortalamaları arasında anlamlı bir farklılık vardır"*, denir.

Denemelerdeki gözlem sayıları eşit olmadığında ise Tukey (1953) ve Kramer (1956) HSD değeri hesaplanırken kullanılan q değerinin

$$q = \frac{Maksimum(\bar{y}_{i.}) - Minimum(\bar{y}_{i.})}{\sqrt{\frac{MS_{Hata}}{2}\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}}, \quad i = 1, 2, \cdots, a \tag{9}$$

olarak değiştirilmesini önermişlerdir, bkz Kuehl (2000). Bu metot Tukey-Kramer metodu olarak da bilinir.

Spjøtvoll & Stoline (1973) ise farklı bir yaklaşımla denemelerdeki gözlem sayıları birbirinden çok fazla farklı olmadığında kritik değer olarak

$$HSD = q_{\alpha}(a, df_{Hata}) \frac{\sqrt{MS_{Hata}}}{\min(\sqrt{n_i}, \sqrt{n_j})} \tag{10}$$

değerinin kullanılmasını önermiştir.

Her bir denemedeki gözlem sayısı birbirinden çok farklı olduğunda ise ilerleyen bölümlerde anlatılacak olan Scheffe metodunun kullanılması önerilmiştir, bkz. Milliken & Johnson (1992).

Duncan çoklu aralık testinde,

$$H_0 : \mu_i = \mu_j, \quad \forall i \neq j \tag{11}$$

hipotezini sınamak için

R_g Test İstatistiği

$$R_g = \begin{cases} r_\alpha(g, df_{Hata})\sqrt{\dfrac{MS_{Hata}}{n}}, & n_1 = n_2 = \cdots = n_a = n \text{ ise} \\ r_\alpha(g, df_{Hata})\sqrt{\dfrac{MS_{Hata}}{a \Big/ \sum\limits_{i=1}^{a}(1/n_i)}}, & \text{en az biri farklı ise} \end{cases} \tag{12}$$

R_g $(g = 2, 3, \cdots, a)$ değerleri hesaplanır.

- Burada $\sum_{i=1}^{a}(1/n_i)$ harmonik ortalamayı ifade eder ve $r_\alpha(g, df_{Hata})$ da Duncan Çoklu Aralık Testinin tablo değerini gösterir, bkz. Duncan (1955).
- Duncan çoklu aralık testinde deneme ortalamalarının ikili karşılaştırmaları aşağıdaki adımlar izlenerek yapılır;

1. **Adım** $\bar{y}_{1\cdot}, \bar{y}_{2\cdot}, \cdots, \bar{y}_{a\cdot}$ deneme ortalamaları

$$\bar{y}_{(1)\cdot} \leq \bar{y}_{(2)\cdot} \leq \cdots \leq \bar{y}_{(a)\cdot}$$

olacak şekilde küçükten büyüğe doğru sıralanır.

2. **Adım** $g = 2, 3, \cdots, a$ için R_g değerleri hesaplanır.

3. **Adım**

$$
\begin{array}{lll}
\bar{y}_{(a)\cdot} - \bar{y}_{(1)\cdot} > R_a & & \\
\bar{y}_{(a)\cdot} - \bar{y}_{(2)\cdot} > R_{a-1} & \bar{y}_{(3)\cdot} - \bar{y}_{(1)\cdot} > R_3 & \\
\vdots & \bar{y}_{(3)\cdot} - \bar{y}_{(2)\cdot} > R_2 & y_{(2)\cdot} - \bar{y}_{(1)\cdot} > R_2 \\
\bar{y}_{(a)\cdot} - \bar{y}_{(a-1)\cdot} > R_2 & &
\end{array}
$$

karşılaştırmaları yapılır.

- Eğer, deneme ortalamaları arasındaki fark R_g değerinden daha büyük ise H_0 hipotezi reddedilir. Bir başka deyişle, "*$i-$inci ve $j-$inci deneme ortalamaları arasında anlamlı bir farklılık vardır*" denir, bkz Montgomery (2001).
- Burada dikkat edilmelidir ki, $\bar{y}_{(a).} - \bar{y}_{(1).} > R_a$ ise $a-$ıncı deneme ile $1-$inci deneme arasında değil, en büyük ortalamaya sahip deneme ile en küçük ortalamaya sahip deneme arasında anlamlı bir fark vardır.

- **Lineer bağıntılar** (linear contrasts) μ_i lerin lineer bileşimleri olarak tanımlanır ve

$$C = \sum_{i=1}^{a} c_i \mu_i \tag{13}$$

olarak gösterilir.

- C nin lineer bağıntı olması için $\sum_{i=1}^{a} c_i = 0$ olması gerekir.

- Lineer bağıntıların sahip olması gereken bir diğer özellik ise **dikliktir** (orthogonality):

$$C_1 = \sum_{i=1}^{a} c_{1i}\mu_{1i} \quad \text{ve} \quad C_2 = \sum_{i=1}^{a} c_{2i}\mu_{2i}$$

iki lineer bağıntı olmak üzere

$$\sum_{i=1}^{a} c_{1i}c_{2i} = 0 \tag{14}$$

eşitliği sağlanıyorsa bu bağıntılar **dik lineer bağıntı** (orthogonal linear contrast) olarak adlandırılır.

- Dikkat edilirse, lineer bağıntılar metodunda ikili karşılaştırmalar da dahil olmak üzere deneme ortalamaları arasındaki olası tüm karşılaştırmalar yapılabilir.
- a tane deneme ortalamasının karşılaştırıldığı bir deneyde en fazla $a-1$ tane dik lineer bağıntı kurulabilir.
- Dolayısıyla, her bir lineer bağıntının serbestlik derecesi "1" dir.
- Lineer bağıntılar yönteminde, hangi deneme ortalamalarının karşılaştırılacağına deney yapılmadan önce karar verilir.

Örneğin, $H_0 : \mu_1 = \mu_2 = \mu_3$ hipotezinin reddedilmesi halinde

$$\begin{array}{lcr} C_1 & : & \mu_1 - \mu_2 = 0 \\ C_2 & : & \mu_1 + \mu_2 - 2\mu_3 = 0 \end{array}$$

şeklinde iki tane lineer bağıntı tanımlanabilir. Burada,

$$c_{11} = 1, \quad c_{12} = -1, \quad c_{13} = 0; \quad c_{21} = 1, \quad c_{22} = 1, \quad c_{23} = -2$$

olduğundan ve

$$\sum_{i=1}^{3} c_{1i} = 0, \quad \sum_{i=1}^{3} c_{2i} = 0, \quad \sum_{i=1}^{3} c_{1i}c_{2i} = 0$$

koşulları sağlandığından C_1 ve C_2 lineer bağıntıları diktir.
Dikkat edilirse, dik lineer bağıntı sayısı denemelerin serbestlik derecesine yani $a - 1 = 2$ ye eşittir.

Lineer bağıntılar yönteminde temel amaç,

$$H_0 : \sum_{i=1}^{a} c_i \mu_i = 0 \tag{15}$$

hipotezini sınamaktır. Bu hipotezi sınamak için

Test İstatistiği

$$F_C = \frac{SS_C \big/ 1}{MS_{Hata}} = \frac{MS_C}{MS_{Hata}} \tag{16}$$

test istatistiği kullanılır.

- Burada,

$$SS_C = n\left(\sum_{i=1}^{a} c_i \bar{y}_{i.}\right)^2 \Bigg/ \sum_{i=1}^{a} c_i^2 \tag{17}$$

dır.

- Eğer,

$$F_C > \boldsymbol{F}_{\alpha;a-1;df_{\boldsymbol{Hata}}}$$

ise H_0 hipotezi reddedilir.

$H_0 : \mu_1 = \mu_2 = \cdots = \mu_a$ hipotezi reddedilmişse ve sınırlı sayıda planlanmış bağıntı karşılaştırılmak isteniyorsa Bonferroni metodunun kullanılması önerilir.

Bonferroni metodu kullanılarak k tane karşılaştırma yapılmak isteniyorsa,

$$H_0 : \sum_{i=1}^{a} c_{ib}\mu_i = 0, \quad b = 1, 2, \cdots, k \tag{18}$$

şeklinde tanımlanan k tane sıfır hipotezinin her biri için ayrı ayrı

Test İstatistiği

$$B_b = t_{\alpha/2k, df_{hata}} \sqrt{MS_{Hata} \sum_{i=1}^{a} \frac{c_{ib}^2}{n_i}}, \quad b = 1, 2, \cdots, k \tag{19}$$

değeri hesaplanır.
Eğer,

$$\left| \sum_{i=1}^{a} c_{ib}\bar{y}_{i\cdot} \right| > B_b$$

ise H_0 hipotezi reddedilir, Milliken & Johnson (1992).

Scheffe (1953) tarafından önerilen ve kendi adıyla anılan metot, Bonferroni metodundan farklı olarak sadece önceden planlanmış az sayıdaki lineer bağıntıyı değil, bütün olası bağıntıları test etmek için önerilmiştir, bkz Kuehl (2000).

Scheffe metodunda

$$H_0 : \sum_{i=1}^{a} c_{ib}\mu_i = 0, \quad b = 1, 2, \cdots, k \tag{20}$$

hipotezini sınamak için

Test İstatistiği

$$S_{C_b} = \sqrt{(a-1)\boldsymbol{F}_{\alpha;a-1;df_{\boldsymbol{Hata}}}}\sqrt{MS_{Hata}\sum_{i=1}^{a}\frac{c_{ib}^2}{n_i}} \tag{21}$$

değeri hesaplanır. Eğer,

$$\left|\sum_{i=1}^{a} c_{ib}\bar{y}_{i.}\right| > S_{C_b}$$

ise H_0 hipotezi reddedilir.